James S. Demetry

OPTIMIZING FOR MINIMUM RESPONSE TIME IN A LINEAR SYSTEM WITH LIMITED CONTROL EFFORT.

UNITED STATES
NAVAL POSTGRADUATE SCHOOL

OPTIMIZING FOR MINIMUM RESPONSE TIME
IN A LINEAR SYSTEM
WITH LIMITED CONTROL EFFORT

By

James S. Demetry

Harold A. Titus

RESEARCH PAPER NO. 45

April 1964

OPTIMIZING FOR MINIMUM RESPONSE TIME IN A LINEAR SYSTEM WITH LIMITED CONTROL EFFORT

By

James S. Demetry

Harold A. Titus

ABSTRACT

Integral indices of performance are brought to bear on the problem of minimizing step response time in some simple control systems. This minimization is subject to a bound on the total control effort that may be expended in response to a step command.

Special attention is given to an index whose integrand is a form of the Euclidean norm of the system's states. An example is given showing, for a given system, the development of a weighting scheme for the second term in a two-term index of performance.

I Introduction

This paper deals with the use of performance indices (integral cost functions) in the design of control systems that are to satisfy certain time-domain requirements and limitations. The optimum design, i.e., minimum index value, is not an end in itself, but a means of obtaining the desired response characteristics with economy of calculation.

II The Optimization Problem

The optimum design of linear, continuous feedback control systems usually implies optimization with respect to some pre-selected index of performance (I P), or cost function. Such indices of performance normally take the form of integrals of various functions of time (t), and error (e). Optimization is possible because the indices generally exhibit minima as functions of a selected system parameter.

As an example of such an optimum design, consider the system of Figure 1, a second-order, type 1 position control system. The value of K is fixed, and the tachometer gain a is to be selected to minimize the index $IP = \int_0^\infty e^2\,dt$ in response to an input $\theta_i = Ru(t)$. By integral tables developed from Parseval's theorem,

$$IP = \int_0^\infty e^2\,dt = \frac{R^2\left[1 + a^2K\right]}{2aK} \qquad (1)$$

Minimizing the IP with respect to a,

$$\frac{\partial(IP)}{\partial a} = R^2\left[\frac{a^2K - 1}{2a^2K}\right] = 0 \qquad (2)$$

we obtain

$$a = \frac{1}{\sqrt{K}} \qquad (3)$$

Using the tachometer gain given by (3) yields, then, a system that is optimum with respect to the integral of square error as an index of performance. Note the familiar result that the damping ratio, ζ, is 0.5.

It should be noted that no practical solution exists for the problem above if it posed so that we seek an optimum value of K for a given a. The solution demands that K be infinite for minimum IP.

Two questions might logically be asked about the analytic design procedure carried out in the preceeding example.

a) Is the minimization of the performance index an end in itself, or is it a means whereby acceptable figures for other, more familiar and perhaps more meaningful measures of system performance are more or less automatically arrived at?

b) Will the design that results from such an optimization always be realizable in the sense that it does not exceed certain energy or equipment limitations?

III A Constraint on Total Control Effort

Question b) will be dealt with first. The particular type of physical limitation to be considered is the constraint of an upper bound on the integral square value of an appropriate quantity within the system. Such a constraint may realistically be interpreted in terms of the total energy expended in driving the system in response to a given input, or perhaps in terms of the heating effect experienced by a component in the system. By way of example, the amplifier output denoted by u in figure 1 will be constrained by the integral form

$$\int_0^\infty u^2 \, dt = \gamma \triangleq \text{total control effort} \qquad (4)$$

The equality sign is intended, rather than $\leq$, since we wish the system to operate at the bound. Integral tables yield, for $\theta_i = Ru(t)$,

$$\int_0^\infty u^2 \, dt = \frac{R^2 K}{2a} = \gamma \tag{5}$$

The immediate result of the constraint is an expression which shows the possible and allowable combinations of amplifier gain K and tachometer gain a; namely,

$$a = \frac{KR^2}{2\gamma} = \frac{K}{2E} \; ; \quad E \triangleq \frac{\gamma}{R^2} \tag{6}$$

IV Optimization Subject to the Constraint

Let us now return to the problem of the optimum design for the system of Figure 1. The minimization of the IP may now be carried out with respect to either parameter, K or a. Note that it is not necessary to pre-specify either of the two gains, since their values are now related by equation (6). This lends a certain measure of uniqueness to the design. Substituting equation (6) into equation (1) gives

$$IP = \frac{R^2 \left[1 + \frac{K^3 R^4}{4\gamma^2} \right]}{K^2 R^2 / \gamma} \tag{7}$$

Minimizing with respect to K;

$$\frac{d(IP)}{dK} = -\frac{2}{K^3} + \frac{1}{4E^2} = 0 \tag{8}$$

which yields

$$K = 2E^{2/3} \tag{9}$$

Calculating the damping ratio for the optimum system, we find

$$2\zeta\omega_n = Ka \tag{10}$$

giving

$$\zeta = \frac{a\sqrt{K}}{2} = \frac{1}{4E} K^{3/2} = \frac{\sqrt{2}}{2} \tag{11}$$

Once γ, the total allowable control effort, is chosen, the optimum design follows from equations (6) and (9). The design may be said to be optimum with respect to integral square error, subject to the constraint on total control effort. That is, choosing a different (non-optimum) set of values for K and a̲ might result in a smaller integral square error, but the upper bound on $\int_0^\infty u^2 \, dt$ will as a result be violated. Note, also, that the optimum design has a damping ratio of 0.707, independent of γ or R.

Introduction of the constraint has not only acknowledged a physical limitation of the plant, but has as well allowed for the simultaneous solution for two optimum parameters within the system.

V A Generalized Performance Index

Answers to question a) have been the subject of many technical papers.[1,2] Gibson says of integral square error as an index of performance, "has poor selectivity. Optimum systems tend to be underdamped. Used mainly because of mathematical convenience." It is for this mathematical convenience that a generalization of the ISE will be considered. The integral square of the Euclidean norm of a system's states is proposed as an index of performance. This index will be used to synthesize minimum-time or quasi-minimum-time linear control in systems where two parameters are variable. The two parameters are

---

[1] Gibson et al, App. and Ind., May, 1961.

[2] Graham and Lathrop, App. and Ind., Nov., 1953.

to be made functionally dependent by an integral square constraint on the signal driving the plant. The minimum time description and the integral square constraint shall apply for the case in which the largest anticipated step input is applied to the system. We seek an algorithm for the selection of the literal coefficients in the generalized index of performance

$$IP = \int_0^\infty \left[x_1^2 + ax_2^2 + bx_3^2 + \dots\right] dt \qquad (12)$$

such that response time is minimized. The coefficients may be functions of R, γ and the plant's fixed parameters; the x_i are the state variables.

VI Definition of Response Time

For purposes of this paper, response time is given a definition in terms of state variables, as follows: T_r is the time required for the state point to move from its initial value $\underline{x}(0)$ to within a certain distance from its desired final value, $\underline{x}(F)$. The state variables may or may not be normalized in determining the distance from the final state.

To illustrate this definition of response time, let us refer to Figure 1 and note the manner in which the state variables x_1 and x_2 are defined. The state trajectory of this system in response to an input $\theta = Ru(t)$ is shown on the phase plane, Figure 2. The response time is the time required for the magnitude of the vector D to decrease from its initial value R, to the value ρ, where ρ is an arbitrary constant value. It is desirable to have the state point remain within this target circle of radius ρ once it has entered, lest the ambiguity of multiple crossings of the target circle be introduced. It is for this reason that the velocity state is best normalized to x_2/ω_n, where ω_n is the undamped natural frequency of the system. It may be shown that such a normalization insures a single crossing of the target circle for any second-order system

responding to a step input.

High-accuracy (digital computer) response time information for the system of Figure 1 is presented in Figure 3. The constraint on total control effort is as given in equations (5) and (6), with γ arbitrarily chosen as 1.0. The independent variable is forward gain K, to each value of which there corresponds a unique value of <u>a</u> by constraint equation (6). Plotted against K is the response time to several target circles on the velocity-normalized phase plane, where the radii of the respective circles are expressed as percentages of the step input magnitude, R. For each of the response time curves there is a setting for K (and hence for <u>a</u>) which minimizes that curve; that setting may be said to be the optimum setting with respect to response time to that <u>particular</u> target circle. It is observed that as the target circle shrinks, response time goes to infinity, with the minimum on the curve moving upward and to the right, asymptotic to a vertical line through K = 2.5. In an absolute sense, then, this system is optimized, with respect to response time to an infinitesimal target circle, by a gain setting of K = 2.5 and the corresponding tachometer setting of <u>a</u> = 1.25.

The information in Figure 3 was, as previously noted, obtained by calculation. Could the same conclusions as to the optimum response time design have been reached by optimizing the system with respect to some mathematically workable index of performance? This, of course, is question a) rephrased, and applied to this example. The answer to it is yes, if there exists an index of performance which, when plotted on Figure 3 with its value versus K, exhibits a minimum coincident

or near coincident with the optimum K for minimum response time in the absolute sense.

VII Constructing the Performance Index

In order to examine the suitability of several common performance indices for obtaining the objective stated above, their values, corresponding to step responses with forward gain set at K, are plotted versus K. The minima of ISE and ITAE, et al, are seen by inspection not to coincide with the optimum K for minimum response time in the absolute sense. That ISE does not provide the IP we seek is not surprising, in view of Gibson's observations. One would expect ITAE to improve on ISE, since time has been added as a penalty factor; and it does, to some extent. It is apparent, however, that functions of time and error alone will not suffice as consistent selectors of optimum parameter settings.

Careful inspection of Figure 2 suggests that we should be concerned not only with functions of x_1 (error) as indices but with functions containing x_2 (velocity) as well, in such a way that the magnitude of vector D is indicated. An index, or cost function, that penalizes duration of D would intuitively seem to hasten the entry of the state point into the target circle. Toward this end, the following index is suggested;

$$IP = \int_0^\infty \left[x_1^2 + \left(\frac{x_2}{\omega_n} \right)^2 \right] dt \qquad (13)$$

where the velocity state has again been normalized by the system's natural frequency. Integral tables are used to evaluate IP as follows:

$$IP = \frac{R^2\left[1 + a^2K\right]}{2aK} + \frac{1}{K}\left[\frac{R^2}{2a}\right] \tag{14}$$

Note that $\omega_n^2 = K$, from the characteristic equation of the system,

$$s^2 + Kas + K = 0 \tag{15}$$

Substituting $a = \frac{K}{2E}$ from equation (6) into equation (14), we have

$$IP = \frac{ER^2\left[2 + \frac{K^3}{4E}\right]}{K^2} \tag{16}$$

For minimum IP,

$$\frac{d(IP)}{dK} = \frac{R^2}{4K^4}\left[3K^4 - 2K(BE + K^3)\right] = 0 \tag{17}$$

from which

$$K = (16E)^{1/3} \tag{18}$$

This yields a damping ratio of

$$\zeta = \frac{a\sqrt{K}}{2} = \frac{K^{3/2}}{4E} = 1.0 \tag{19}$$

We note that the optimum damping ratio of unity is independent of γ or R. In the case for which Figure 3 was drawn, $\gamma = R = 1.0$, giving an optimum gain K, by equation (18), of 2.52. This result compares very well with the calculated responses as presented in Figure 3. An index of performance has been found whose minimum coincides with that of another performance characteristic; namely, minimum response time in the absolute sense, in the presence of an energy constraint.

VIII Another Example

Let us now apply the IP of equation (13) and the constraint of

equation (4) to another second-order system, that of Figure 4, in which the plant has a non-zero pole. For step inputs $\theta_i = Ru(t)$, integral tables yield

$$\int_0^\infty u^2 dt = \frac{R^2K\left[K + p^2\right]}{2(p + aK)} = \gamma \tag{20}$$

The IP of equation (13) is

$$IP = \frac{R^2K + R^2(p + Ka)^2}{2K(p + Ka)} + \frac{1}{K}\left[\frac{R^2K^2}{2K(p + Ka)}\right] \tag{21}$$

This reduces to

$$IP = \frac{2R^2K + R^2(p + Ka)^2}{2K(p + Ka)} \tag{22}$$

Equation 20 is then solved for $(p + Ka)$, the result being used to eliminate <u>a</u> from equation (22). The IP then becomes

$$IP = \frac{R^2}{4E}\frac{(8E^2 + K^3 + 2K^2p^2 + Kp^4)}{(K^2 + Kp^2)}; \quad E = \frac{\gamma}{R^2} \tag{23}$$

Differentiation for minimization yields the fourth-degree polynomial in K,

$$K^4 + 2p^2K^3 + p^4K^2 - E^2(16K + 8p^2) = 0 \tag{24}$$

Unfortunately, a direct literal expression for K in terms of E and p is not obtainable. The equation may be solved for given values of p and E, however, and the results presented graphically, as they are in Figure 5.

We have yet to show, for this example, that the index of performance minimizes at a value of K such that the resulting system is optimum or near optimum with respect to response time. To show this, we must test for the damping ratio associated with each solution of equation (24).

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

Solving for ζ from the characteristic equation, we have

$$\zeta = \frac{p + Ka}{2\sqrt{K}} \tag{25}$$

where K is obtained from (24) and a from (20). We may generalize the results of the previous example to state that the time-optimal response is characterized by a ζ of 1.0 for these second-order systems. The contours of constant ζ superimposed on Figure 5 give an indication of how near or far the index used comes to selecting a time-optimal design.

Figure 5 reveals that the IP of equation (13) is not a consistent selector of optimum parameters; i.e., the damping ratio ζ of the selected system is dependent upon the pole position of the plant, and upon the total control effort allowed for the response to a step input. This shortcoming may be overcome, however, by taking advantage of the inherent flexibility of a multi-term index of performance. We notice that the velocity term is normalized by the natural frequency, ω_n, but no further weighting has yet been indicated. A weighting factor β is now introduced such that

$$IP = \int_0^\infty \left[x_1^2 + \beta\left(\frac{x_2}{\omega_n}\right)^2\right] dt \tag{26}$$

It is postulated that for any p and R in the problem of Figure 4, there exists a β such that the IP of equation (26), subject to the constraint of equation (20), selects parameters K and a to yield a damping ratio of unity. Minimization of equation (26) subject to (20) gives

$$\beta = \frac{K^2\left[K + p^2\right]^2 - 4E\left[2K + p^2\right]}{4E^2\left[2K + p^2\right]} \tag{27}$$

where K is determined by the simultaneous solution of the constraint equation (20) and the damping equation (25) with $\zeta = 1.0$. The solutions of equation (27) are presented in Figure 6.

Figure 6 constitutes, in graphical form, an algorithm for the selection of the first weighting factor in the general IP of equation (12), as it applies to the minimization of response time in the system of Figure 4. It is seen that the weighting factor β will lie between $\pm$ 1.0. It may be shown that the IP itself remains positive, and that the effect of the β factor is one of shifting the minimum of the IP to coincide with a desired optimum design.

IX A Third-Order Case

It is possible, of course, to claim that the design problems of the preceeding examples could have been carried out without the use of the performance index. For example, the constraint equation (20) and the damping ratio equation (25) with $\zeta = 1$ can be solved simultaneously to give a unique solution that is time optimal. For higher-order systems, however, the significance of ζ and its relationship to response time effectively vanish, leaving us without sufficient equations to solve the optimization problem. It is in the realm of higher-order systems, then, that analytical design by cost functions of the form of equation (10) holds some promise.

Consider, for example, the system of Figure 7. The values of

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

K and $\underline{a}$ are to be selected to minimize response time, subject to the constraint on total control effort,

$$\int_0^\infty u^2 dt = \frac{R^2 K(Ka + p^2)}{2(pa - 1)} = \gamma \tag{28}$$

For $\gamma/R^2 = 10$ and $p = 1$,

$$a = \frac{K + 20}{20 - K^2} \tag{29}$$

The following IP is suggested:

$$IP = \int_0^\infty \left[x_1^2 + \frac{1}{K} x_2^2 + \frac{1}{K^2} x_3^2 \right] dt \tag{30}$$

The selection of the normalizing factors follows the rationale that higher loop gain will give rise to greater magnitudes in the derivative states. Note that no attempt has been made to otherwise weight the factors in the IP. Evaluation of equation (30) yields

$$IP = \frac{R^2 \left[3 - Ka + Ka^2 p \right]}{2K \left[a - 1 \right]} \tag{31}$$

The index may either be minimized analytically, as has been done in the previous examples, or equations (31) and (29) may be solved as K is allowed to vary over a region within which the minimum value of IP lies. In either case, the labor involved is considerably less than that entailed in the design by repeated solution for the response of the system.

Figure 8 shows the time-to-target-sphere* and IP information for

---

* The target sphere is measured on x_1, $\frac{x_2}{K}$, and $\frac{x_3}{K^2}$ coordinates.

the problem of Figure 7. ITAE is also included for comparison. Note that the IP of equation (30) does fairly well in locating its minimum near the time-optimal system. It is entirely plausible that the weighting factor argument of the previous example could be extended to this problem, although this has not been done.

X Conclusions

It has been shown that multi-term indices of performance offer, through proper weighting of their several terms, a means of selecting system parameters to optimize a linear system with respect to response time. It is conceivable that performance characteristics other than response time may be optimized in like manner. No direct and general process for the selection of a weighting scheme has yet emerged from the studies made to date. Further investigation is being carried on into weighting schemes, constrained and unconstrained systems, translation of time and frequency domain specifications into properly selective performance indices, and other topics related to the direct analytic design of control systems.

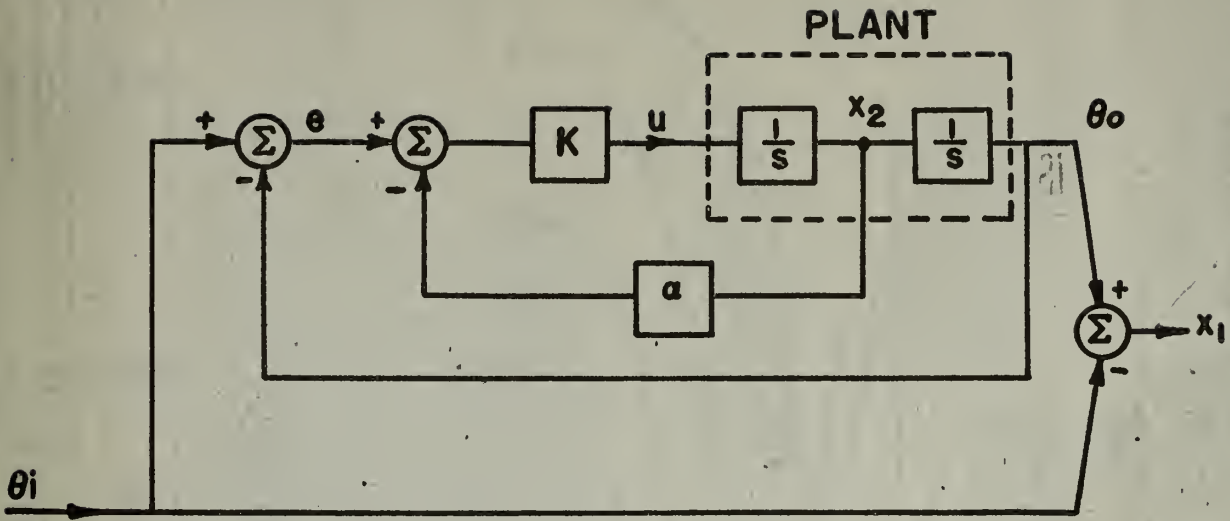


FIGURE 1
A POSITION CONTROL SYSTEM

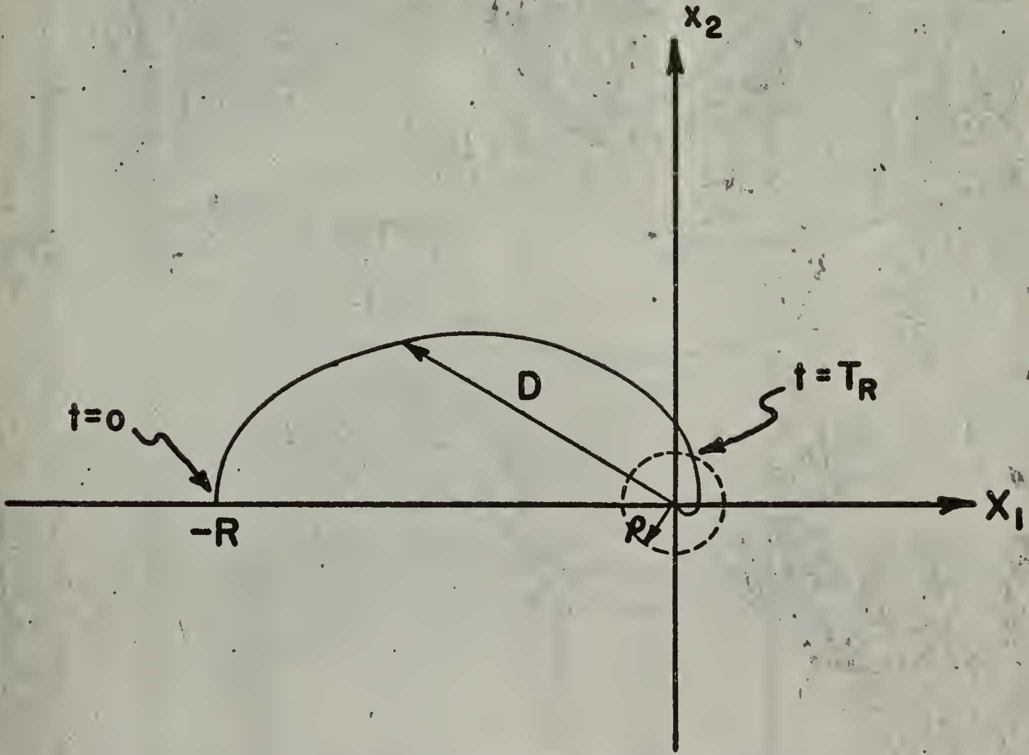


FIGURE 2
STATE TRAJECTORY FOR $\theta i = R\,u(t)$

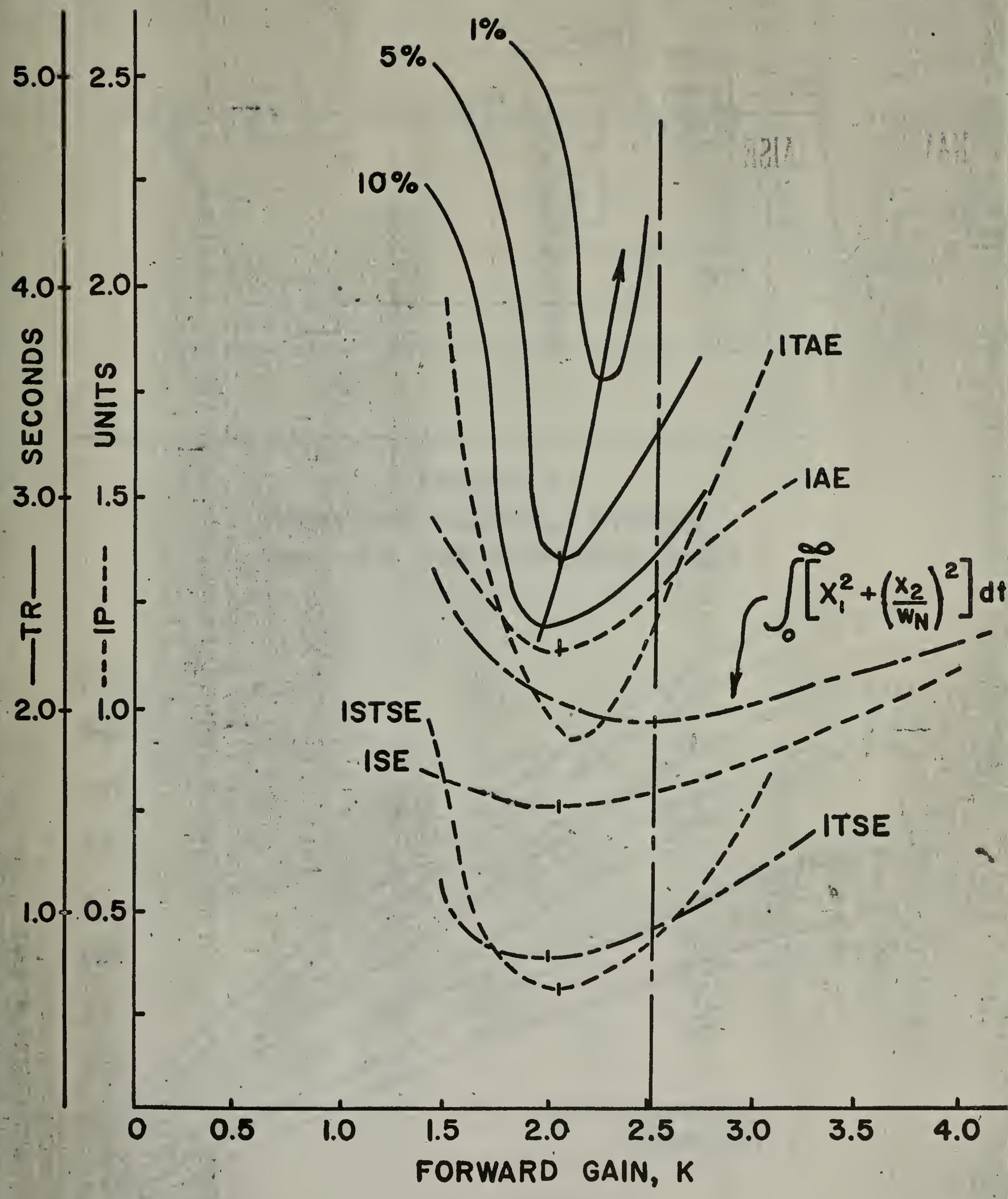


FIGURE 3

RESPONSE TIME AND PERFORMANCE INDICES AS FUNCTIONS OF GAIN K FOR THE SYSTEM OF FIGURE 1, CONSTRAINED

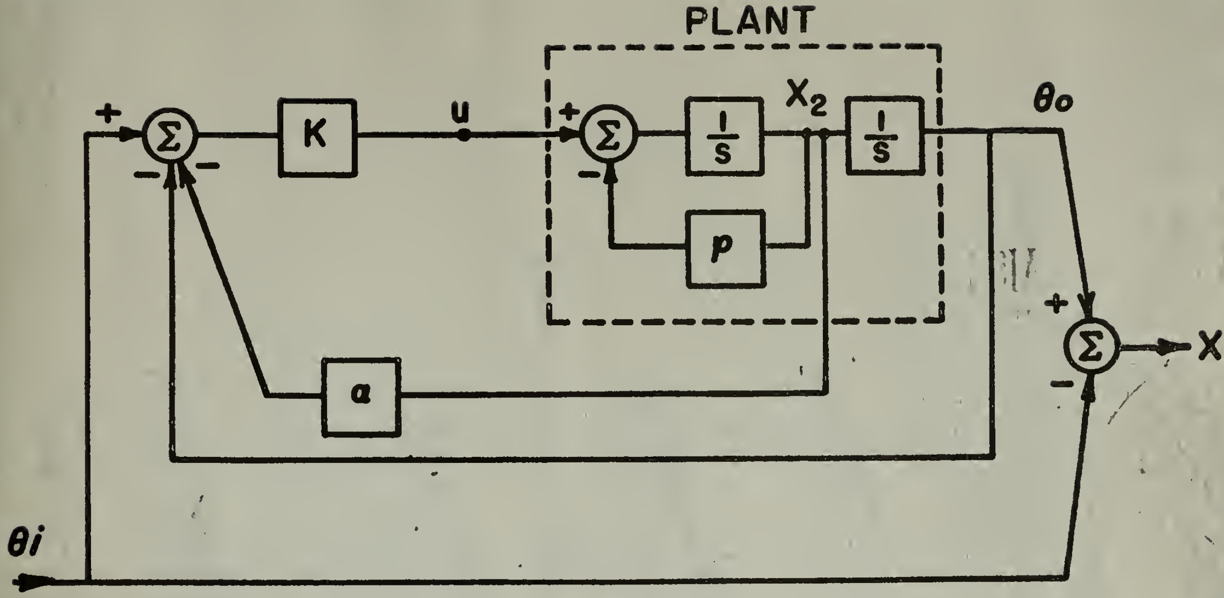


FIGURE 4
POSITION CONTROL SYSTEM
PLANT HAS ONE NON-ZERO POLE

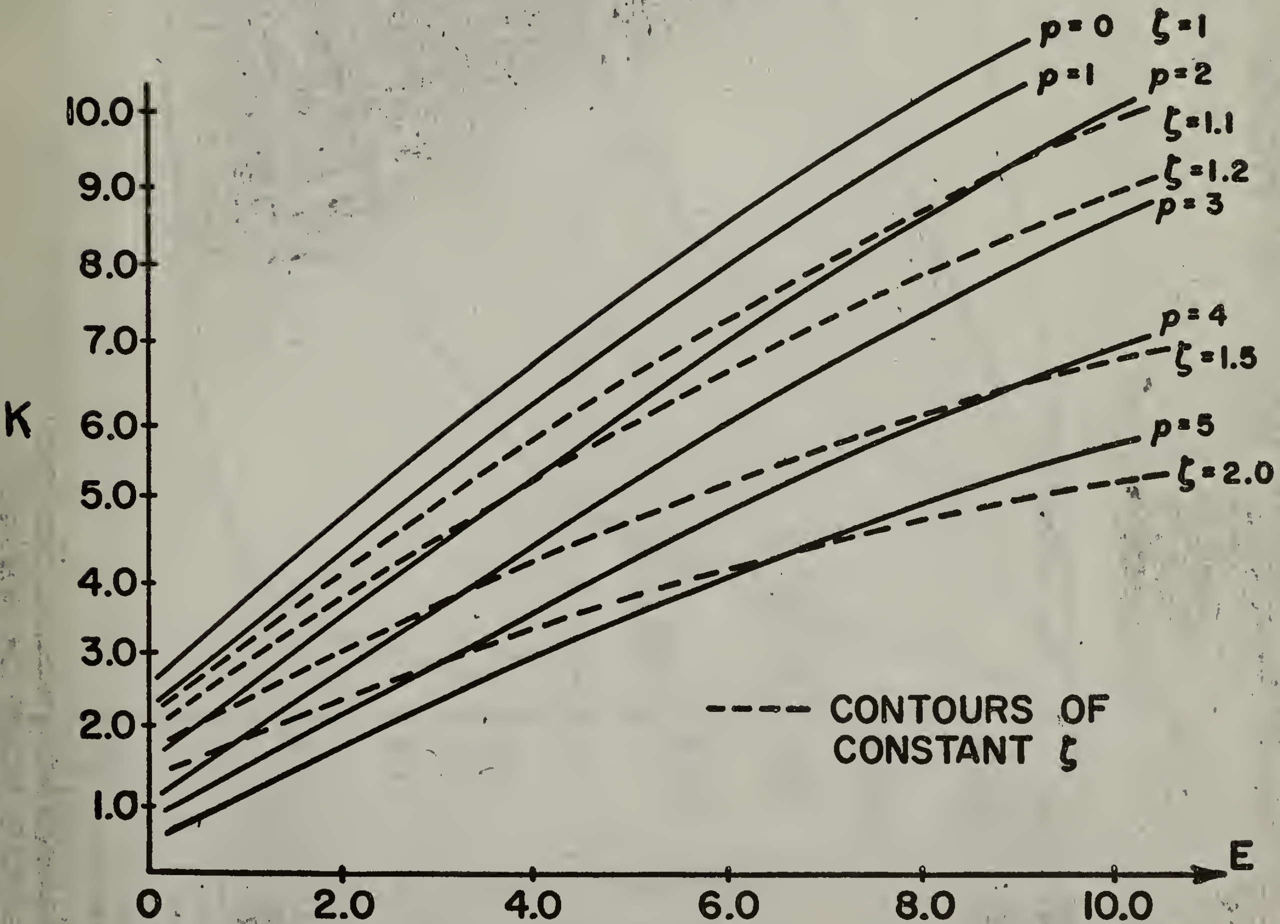
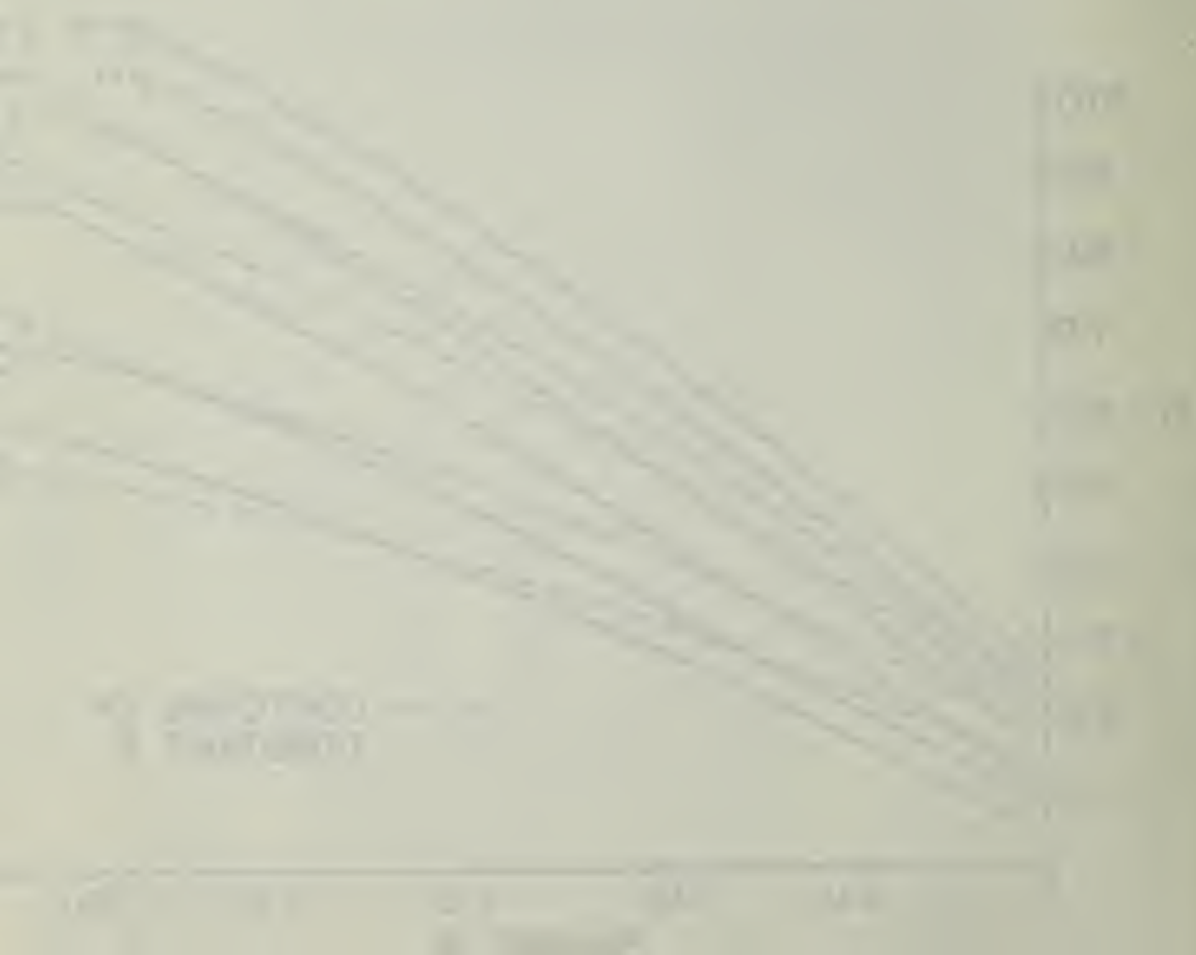


FIGURE 5
SOLUTIONS TO EQUATION 24



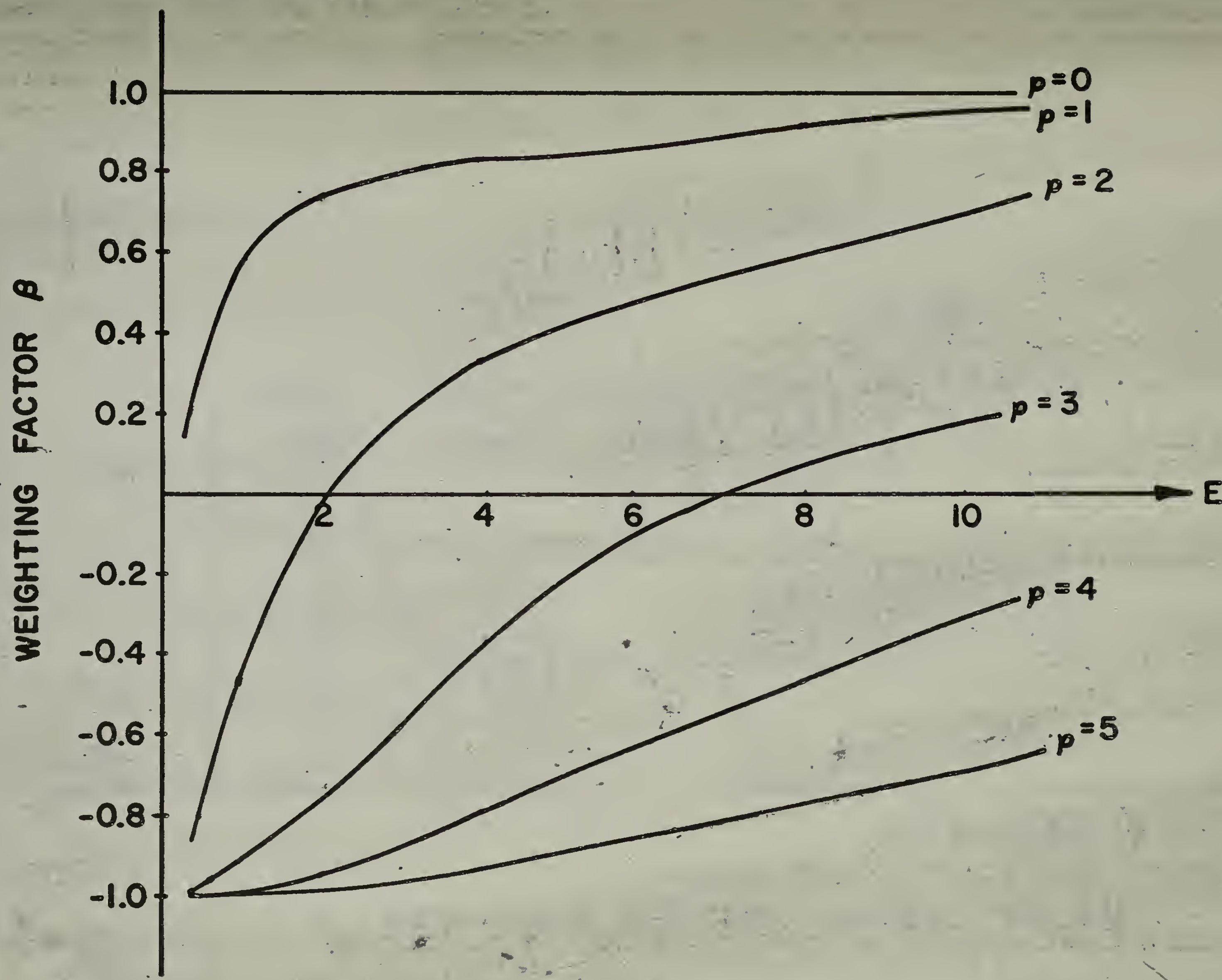


FIGURE 6

WEIGHTING FACTOR FROM EQUATION 27

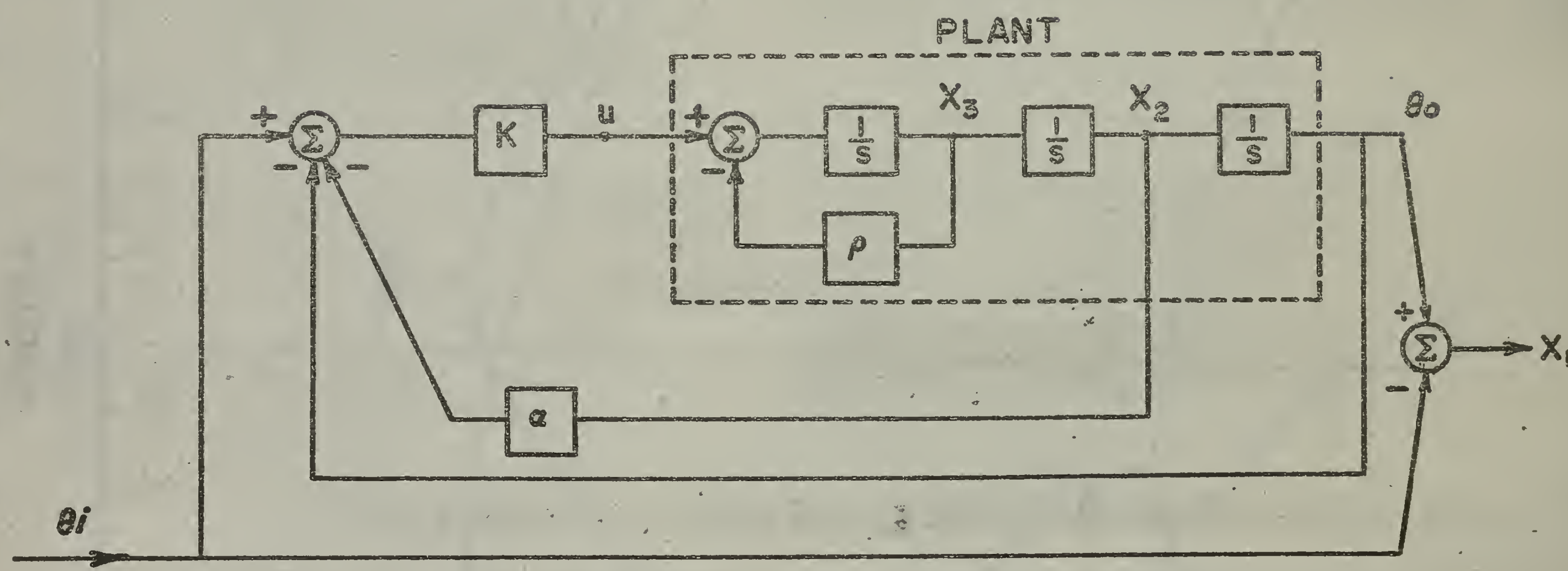


FIGURE 7
A THIRD ORDER POSITION CONTROL SYSTEM

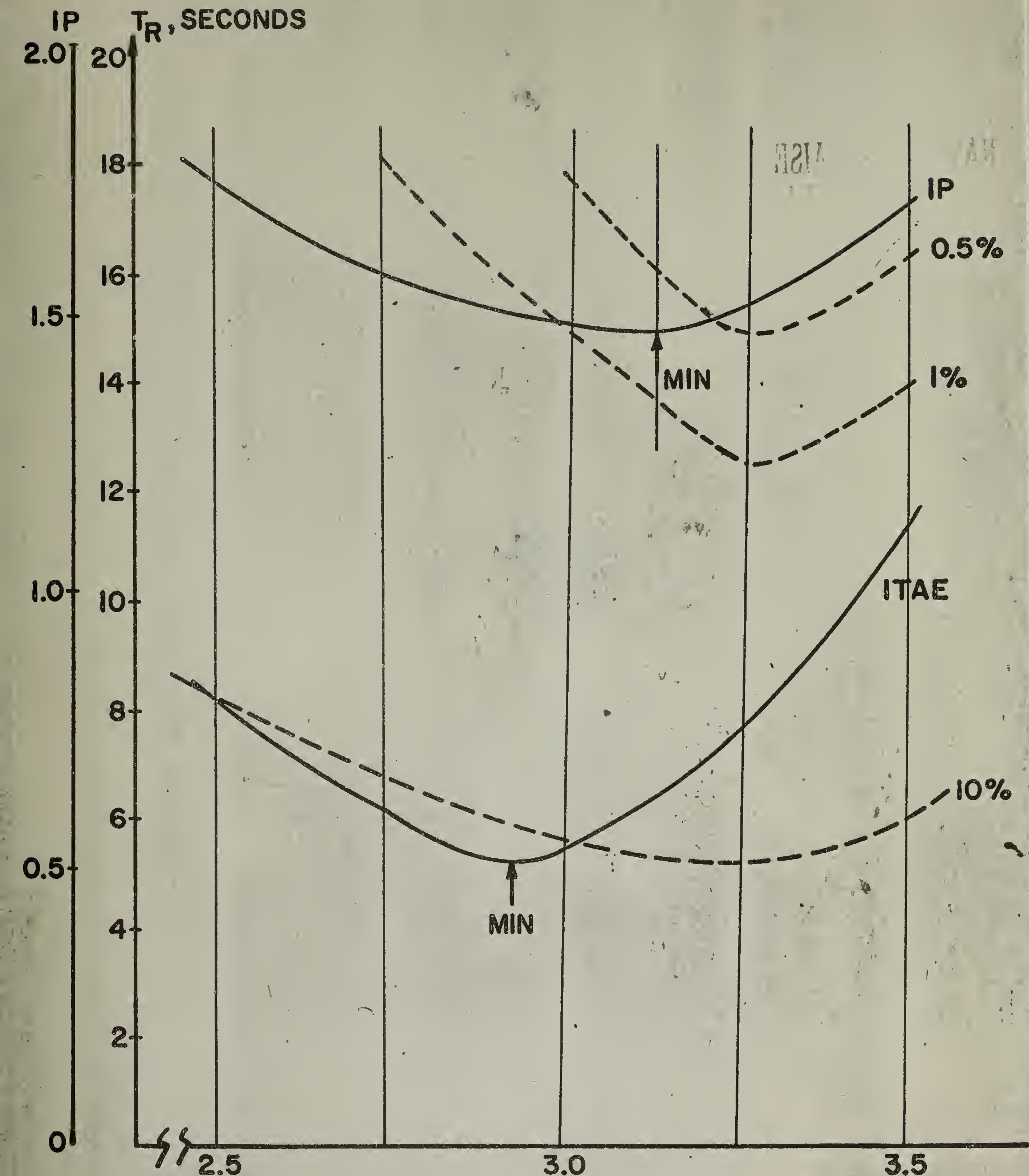


FIGURE 8

RESPONSE TIME AND PERFORMANCE INDICES FOR SYSTEM OF FIGURE 7